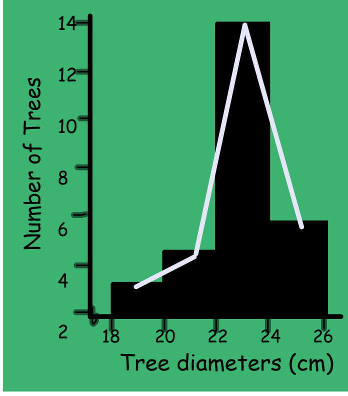
Section 1.5 - Large Distributions and the Normal Curve

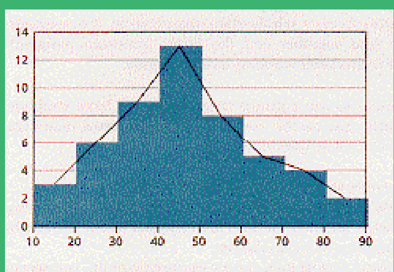
Curriculum Outcomes	Related Activities	Page in Text
• solve problems sing graphing technology	• investigate the impact of increasing the sample size on the shape of the distribution (frequency polygon)	33
• analyze statistical summaries, draw conclusions, and communicate the results about distributions of data • calculate various statistics using appropriate technology, analyze and interpret the displays, and describe relationships	• investigate the impact of graphing data using smaller bin sizes on the appearance of the frequency polygon	34
• explore measurement issues using the normal curve	• use the standard deviation to determine the range within which 68% and 95% of the population falls in a normal distribution	36
	• use standard deviation to identify outliers in a normal distribution	36

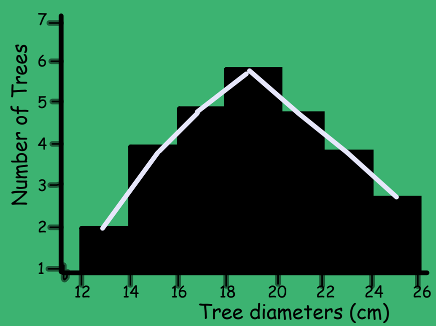
- When large samples of data are chosen at random and are created into a histogram plot they result in a characteristically symmetrical shape.

Frequency polygon = the shape that is formed when the centres of the tops of the bars of a histogram are joined by straight lines.

Examples:







Complete Invest. #4

Investigation #4
Page 33

xmin=150
xmax=200
xscl=10 (bin size)
Ymin=0
Ymax=10
Yscl=2
Xres=1

Purpose

To compare the distribution of a small set of data values to the distribution of a larger set.

Procedure

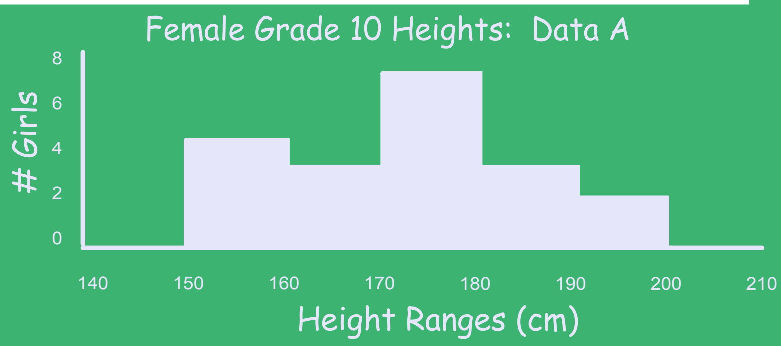
- Create a histogram using a bin width of 10 for this set of height measurements (in centimetres) taken from a **random** sample of 19 Grade 10 girls.

Set 1

150	156	158	159	165	165	169	170	170	172
173	174	175	176	185	186	186	192	193	

- Find the mean, median and mode for each set of data
- Create a frequency polygon for each histogram.

# data = 19
mean = 172
Median = 172
Mode = 165, 170 and 186



xmin=150
xmax=200
xscl=10 (bin size)
Ymin=0
Ymax=35
Yscl=5
Xres=1

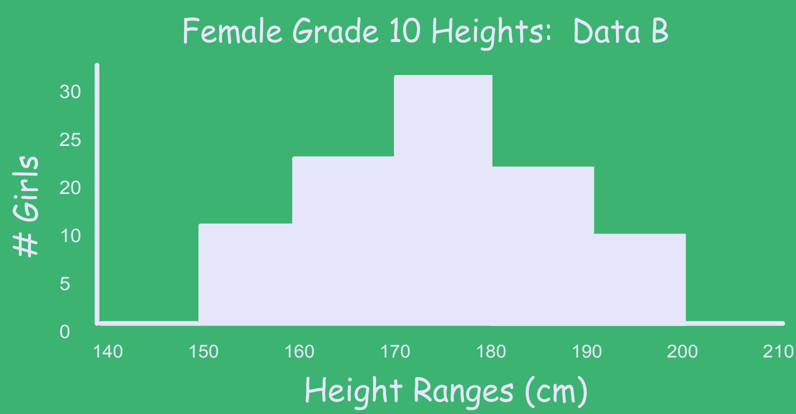
- Create a histogram using a bin width of 10 for this set of height measurements (in centimetres) taken from a **random** sample of 89 Grade 10 girls.

Set 2

150	152	153	154	154	155	156	157	157	160	160	161	161
162	164	164	165	165	166	166	166	167	167	168	168	168
168	169	169	169	169	170	170	170	170	170	171	171	171
171	171	172	173	174	174	174	174	174	174	174	175	176
176	176	176	177	178	178	178	179	179	180	181	182	182
182	182	183	183	185	185	186	186	186	186	186	187	187
188	188	189	192	193	193	194	194	195	195	195		

- Find the mean, median and mode for each set of data
- Create a frequency polygon for each histogram.

# data = 89
mean = 174
Median = 174
Mode = 174



D. Draw vertical lines on the graph at each value.

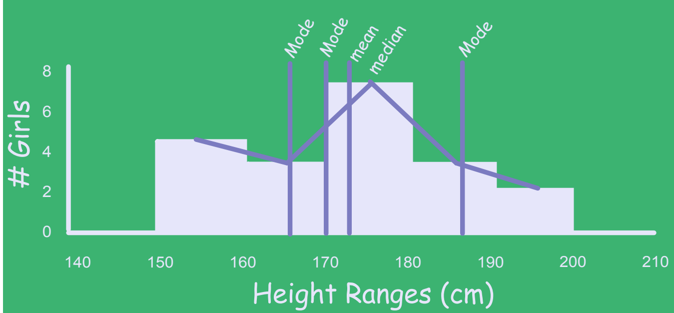
E. Describe: **What do you think??**

- how the shape of Set 1 compares to the shape of set 2
- how the locations of the averages in both sets compare.

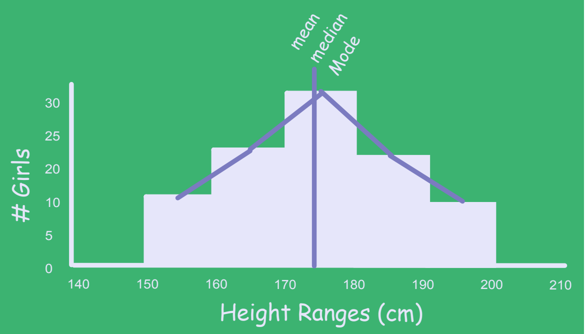
# data = 19
mean = 172
Median = 172
Mode = 165, 170 and 186

# data = 89
mean = 174
Median = 174
Mode = 174

Female Grade 10 Heights: Data A

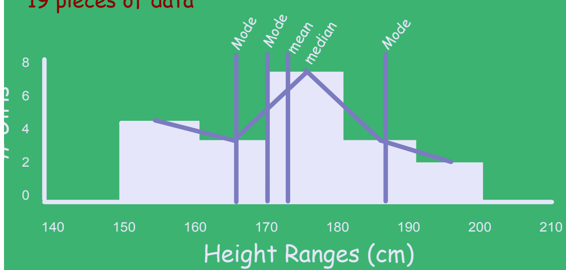


Female Grade 10 Heights: Data B



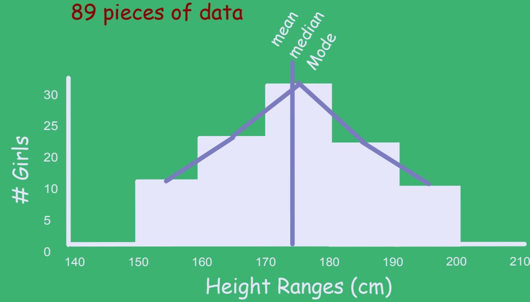
Female Grade 10 Heights: Data A

19 pieces of data



Female Grade 10 Heights: Data B

89 pieces of data



Questions:

1. How does the shape of the histogram from Data A compare to the shape of the histogram from Data B?
2. How do the locations of the averages in both sets of data compare?
3. Suppose you are the manufacturing manager for a clothing outlet. You want to measure the height of teenage girls to find an "average" height. Why would you expect:

a) a larger set of height measurements to have more values in the middle and fewer at the extremes than a smaller set of data?

b) the mean, median, and mode to be closer to each other in a set of 200 measurements than they would be in a smaller set of data?

Conclusions:

If there is a large number of randomly chosen data values it is likely that:

- the data are symmetrical around the middle
- the 3 measures of central tendency are close to the middle
- most data is clustered in the middle
- are few extreme values

Classwork/Homework:

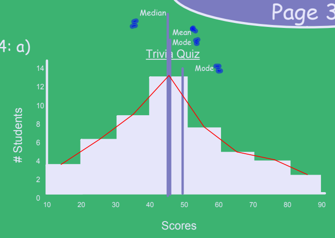
Complete the following on page 35:

- Question # 4 (all out by hand)
- Question #3

Answers:

Page 35

#4: a)



Mean = 45.86
Median = 45
Mode = 45, 50

b) - most data at median, mode
- less data found below mean
- More data found above the mean

c) averages would be found in the middle

d) $\sigma x = 17.61$

$\bar{x} = 45.86$

1 s.d.

45.86 ± 17.61

$= 27.4 - 62.6$

$\rightarrow \frac{32}{50} = 64\%$

2 s.d.

$9.78 - 80.22$

$48/50 \Rightarrow 96\%$

Math 10 - October 21st

- Reminder:

=> If you were absent yesterday, then you need to come in at lunch to write the test.

=> Unit 1 Test on Tuesday, October 26th

- Notes on Normal Distribution
- Classwork/Homework

Normal Distribution

-->When some large pieces of data (10,000 pieces) are plotted in a histogram we can begin to see the resemblance to a bell shape.

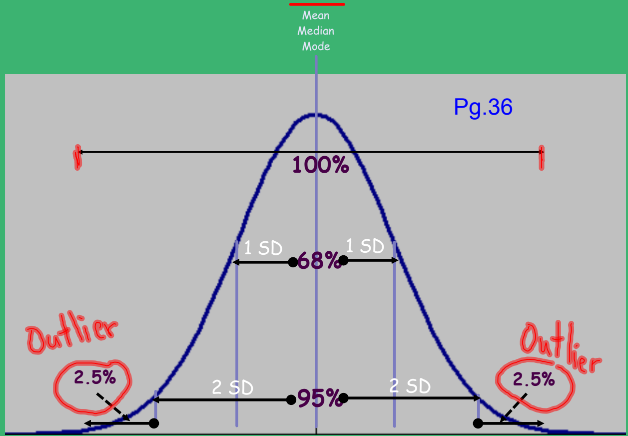
--> This bell curve is known as a Normal Curve and the curve's distribution is known as Normal Distribution.

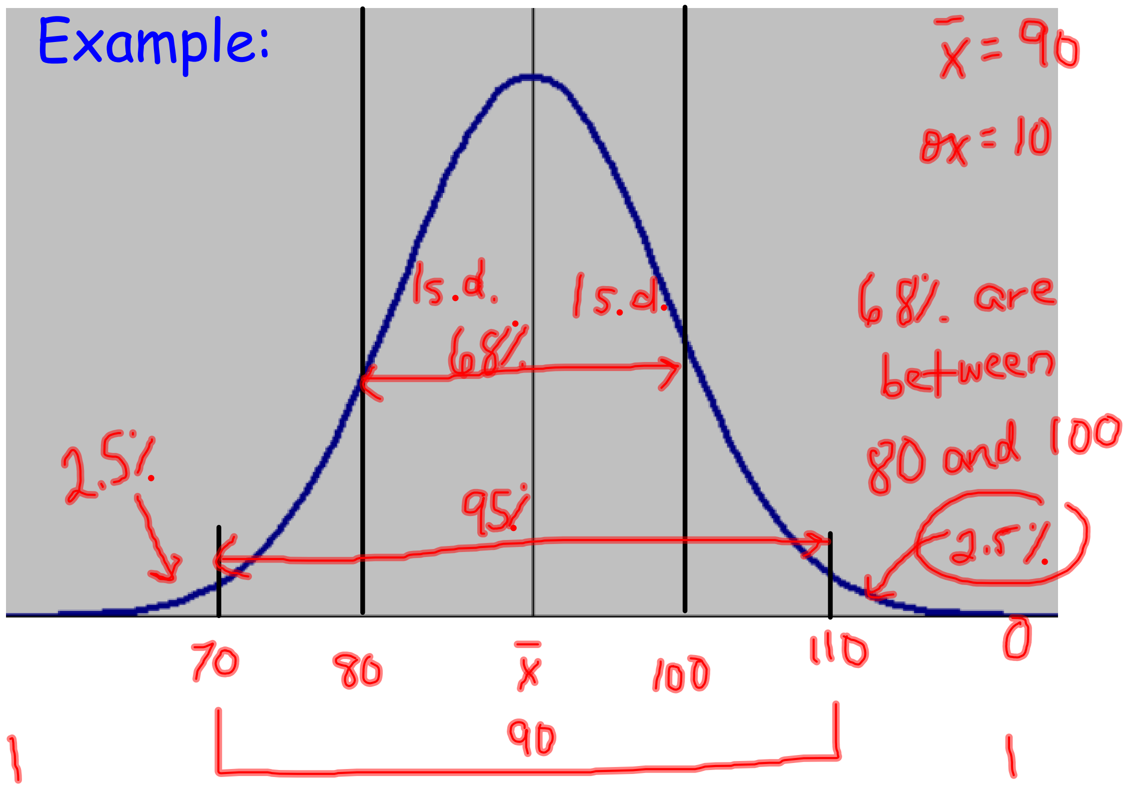
--> Some examples of normally distributed data are:
heights of 15 year olds, pitching speeds of major league pitchers, etc...

--> Some data, no matter how large or random the sample is, will not be normally distributed. An example being rolling a dice many times and record the frequency of each outcome.

For Normal Distribution to occur:

- the shape of the polygon will appear to be a continuous bell-shaped curve:
- the mean, median and mode will all have the same value and be in the middle of the graph.
- the graph of the data will be symmetrical about the middle.
- 68% of the data will be within 1 σx of the mean and 95% of the data will be within 2 σx of the mean.





68% $\Rightarrow \bar{x} \pm \sigma x$

mean ± one standard dev.

95% $\Rightarrow \bar{x} \pm 2\sigma x$

Pg. 37 #5

$\bar{x} = 115$

$\sigma x = 25$

a) 68% 115 ± 25

90 to 140

b) 95% (Most) 115 ± 50

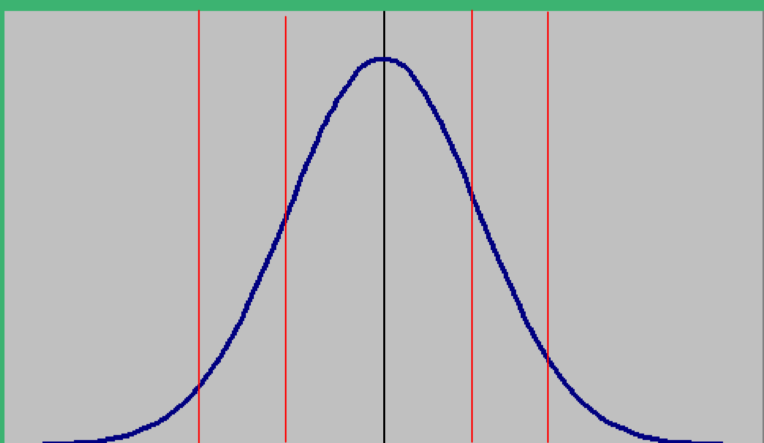
$\frac{2 \text{ s.d.}}{25 \times 2}$

65 to 165

Normal Distribution Curve:

Example #1:

- When a large, random sample of the population's IQ's are recorded in a histogram it will show a normal distribution curve.
- Construct a normally distributed curve of the IQ's. The mean of the IQ's is 110 and the SD is 15.



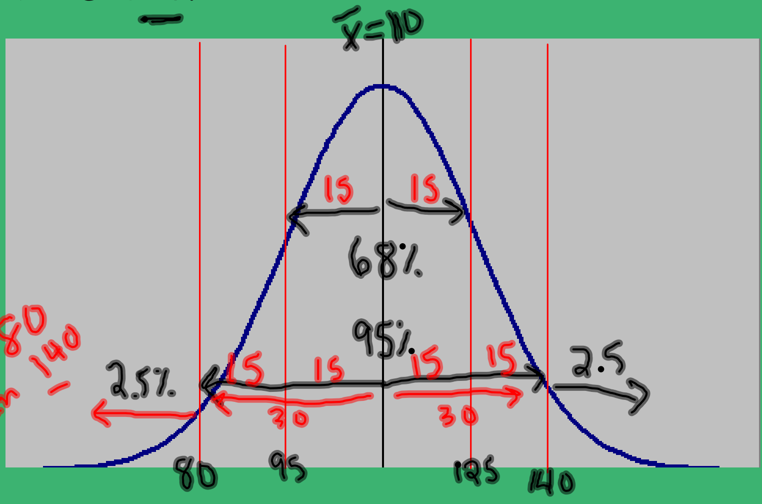
Example #1:

- When a large, random sample of the population's IQ's are recorded in a histogram it will show a normal distribution curve.
- Construct a normally distributed curve of the IQ's. The mean of the IQ's is 110 and the SD is 15.

68%
95 and 125

95%
80 and 140

2.5
less than 80
more than 140

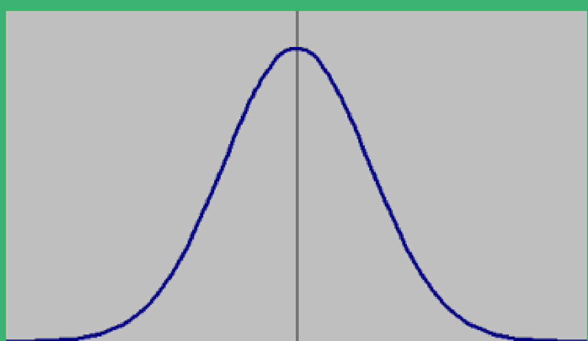


Focus Questions:
Page 37

5. Suppose 960 students took a math-achievement test out of 175.
- The results were normally distributed
- The mean was 115
- The SD was 25

a) Within what range of scores would 68% of the students typically fall? Explain.

b) Within what range would most students fall? Explain?



Focus Questions: Page 37

5. Suppose ~~960~~ students took a math-achievement test out of 175.

- The results were normally distributed
- The mean was 115
- The SD was 25 $\times 2 = 50$

1 s.d.

a) Within what range of scores would ~~68%~~ of the students typically fall? Explain.

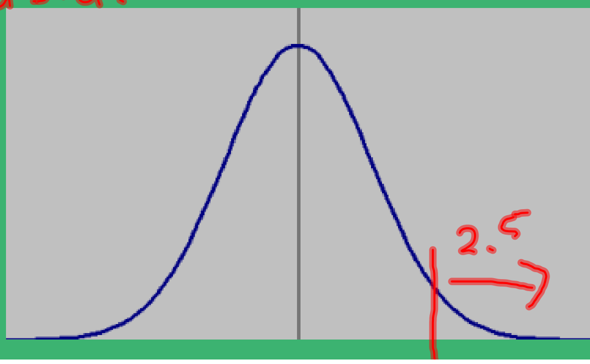
$115 \pm 25 = 90$ to 140

b) Within what range would most students fall? Explain?

95% → 2 s.d.

115 ± 50

$= 65$ to 165



pg. 37-38

#9, 10, 12, 13